Automated Wind Turbine Pitch Fault Prognosis using ANFIS

Bindi Chen
Durham University, UK
bindi.chen@durham.ac.uk

Peter C. Matthews
Durham University, UK
p.c.matthews@durham.ac.uk

Peter J. Tavner
Durham University, UK
peter.tavner@durham.ac.uk

Abstract

Many current wind turbine (WT) studies focus on improving their reliability and reducing the cost of energy, particularly when WTs are operated offshore. WT Supervisory Control and Data Acquisition (SCADA) systems contain alarms and signals that provide significant important information. A possible WT fault can be detected through a rigorous analysis of the SCADA data. This paper proposes a new method for analysing WT SCADA data by using Adaptive Neuro-Fuzzy Inference System (ANFIS) with the aim to achieve automated detection of significant pitch faults. Two existing statistical analysis approaches were applied to detect common pitch fault symptoms. Based on the findings, an ANFIS Diagnosis Procedure was proposed and trained. The trained system was then applied in a wind farm containing 26 WTs to show its prognosis ability for pitch faults. The result was compared to a SCADA Alarms approach and the comparison has demonstrated that the ANFIS approach gives prognostic warning of pitch faults ahead of pitch alarms. Finally, a Confusion Matrix analysis was made to show the accuracy of the proposed approach.

Keywords: Wind Turbine, SCADA, Neuro-Fuzzy, ANFIS, Fault Prognosis, Fault Detection.

1 Introduction

Operation and Maintenance (O&M) costs constitute a sizeable share of the annual cost of a wind farm (WF) and turbine downtime. According to [1][2], the percentage of O&M cost of some European WFs are 12% for onshore and 23% for offshore. With an annual average growth of 15.6% over the last 17 years and more offshore WTs will be deployed in the near future within EU [3], there is a large commercial interest in the economical operation and increased reliability of wind turbines.

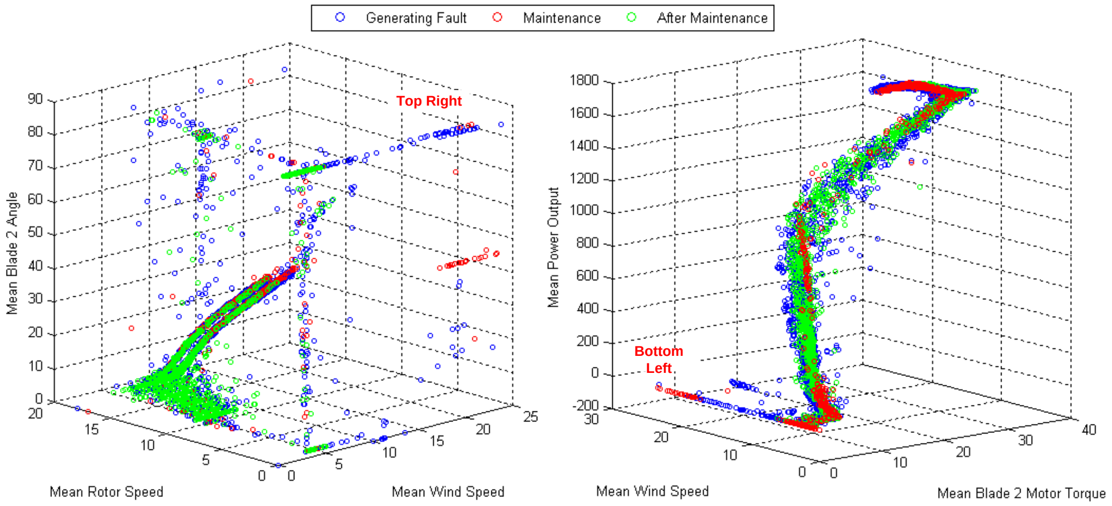
The essence of improving the WT reliability is to reduce the downtime and increase its availability by optimising both the WT design and the maintenance schedule [4]. Both these strategies require a full understanding of the WT system and a detailed analysis of its failure mechanisms. WT SCADA systems provide a rich resource to achieve this capability as it archives comprehensive signal information, historical alarms and detailed fault logs, as well as environmental and operational conditions. A WT's systematic performance can be monitored through a rigorous analysis of the information collected by the SCADA system which covers all the major WT sub-assemblies.

Studies using SCADA data to detect WT faults have been researched during the last 4 years [5]. Some of the more recent methods include signal-based analysis approaches for WT gearbox and generator [6], a system called SIMAP based on artificial neural network aimed to detect and diagnose gearbox failures [7], a probability analysis of pitch performance curves for identifying faults in pitch system [8], an automated analysis system also based on artificial neural network [9], time-sequence and probability-based analysis method to rationalise and reduce SCADA alarm data [11], a pattern recognition approach for identifying faults in WT pitch system [10], and a further study of Venn Diagram analysis using Bayesian Network [12].

It can be seen from above literature that SCADA data volume is usually too large and alarm information is too unclear to indicate failure root cause. This highlights the need for more intelligent methods that can use existing

WF	WT	Case	Generating Fault	Maintenance	After Maintenance
1	A	Case 1	05/01/2008 ~ 15/02/2008	16/02/2008 ~ 21/02/2008	22/02/2008 ~ 03/03/2008
		Case 2	20/12/2006 ~ 14/01/2007	15/01/2007 ~ 25/01/2007	26/02/2007 ~ 10/02/2007
	B	Case 3	22/08/2007 ~ 04/09/2007	05/09/2007 ~ 09/09/2007	10/09/2007 ~ 18/09/2007
		Case 4	17/10/2006 ~ 28/10/2006	29/10/2006 ~ 29/10/2006	30/10/2006 ~ 04/11/2006
		Case 5	10/08/2008 ~ 27/08/2008	28/08/2008 ~ 30/08/2008	31/08/2008 ~ 10/09/2008
		Case 6	20/09/2006 ~ 13/10/2006	14/10/2006 ~ 19/10/2006	19/10/2006 ~ 22/10/2006

Table 1: Six pitch fault cases.



(a) (b)

Figure 1**:** (a) Typical variable-speed pitch-to-feather control plot for Case 1; (b) Pitch torque power curve plot for Case 1;

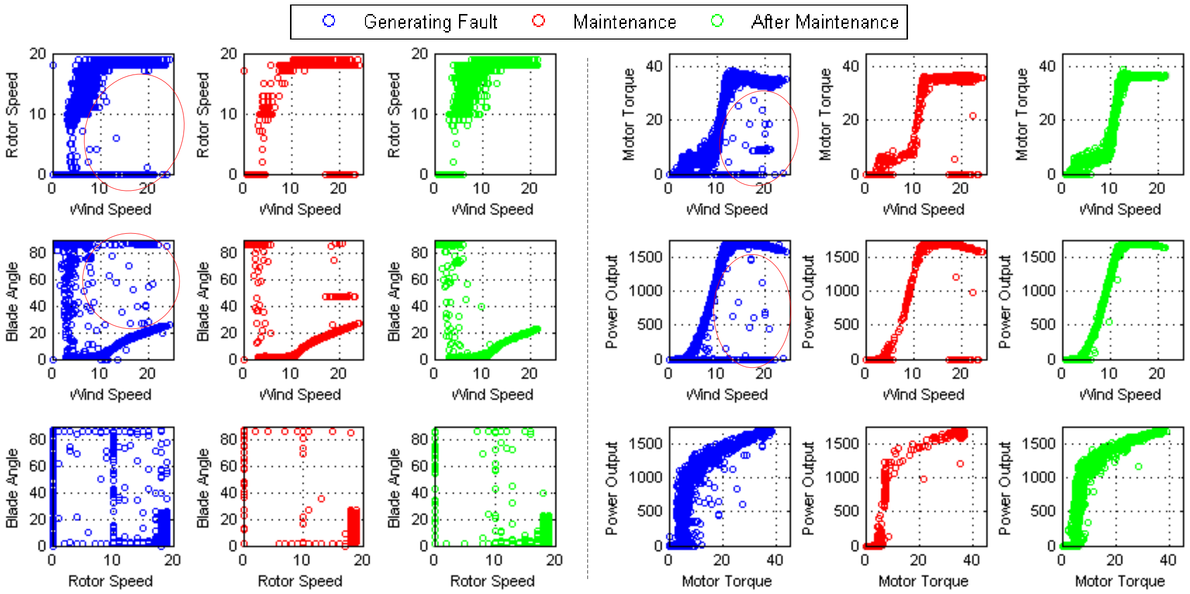
SCADA data to automatically provide accurate WT failure diagnosis. This paper proposes a new method for analysing WT SCADA data by using ANFIS with the aim to achieve automated detection of significant pitch faults.

2 Pitch Fault Analysis

A statistical analysis of six known pitch faults (Cases 1-6, as shown in Table 1), using both typical variable-speed pitch-to-feather control strategy [13] and pitch torque power curve characteristics [8] obtained from SCADA data and Maintenance Records, has been made to find the common pitch fault symptom, as shown in Figure 1.

In Figure 1(a), no data can be found on top right corner in *After Maintenance* period. A normal running turbine should not feather its blade and have zero rotor speed when wind speed is larger than cut-in. Thus, any data appearing on top right corner of this 3D plot can be regarded as a possible pitch fault. For Figure 1(b), no data can be found on bottom left corner in *After Maintenance* period. This is because a normal running turbine should start generating power when the wind speed is greater than cut-in. Meanwhile, blade pitch motor torque is needed to change the blade angle to avoid rotor overspeed. Thus, any data appearing in the bottom left of this 3D plot could be caused by a pitch fault. Although presented in three dimensions, analysis in each of the planes simplifies algorithm development to two variables, 2D views are shown in Figure 2. Then, four 2D views, known as features, were found and can be used to identify wind turbine pitch faults, as encircled in Figure 2.

In addition, a day-by-day analysis of the *Generating Fault* period checking against SCADA alarms had shown that the SCADA signals are able to provide fault detection and much earlier than SCADA alarms, as shown in Figure 3.



(a) Corresponding to Fig.1(a) (b) Corresponding to Fig.1(b)

Figure 2: 2D views of Fig. 1 covering *Generating Fault*, *Maintenance* and *After Maintenance*

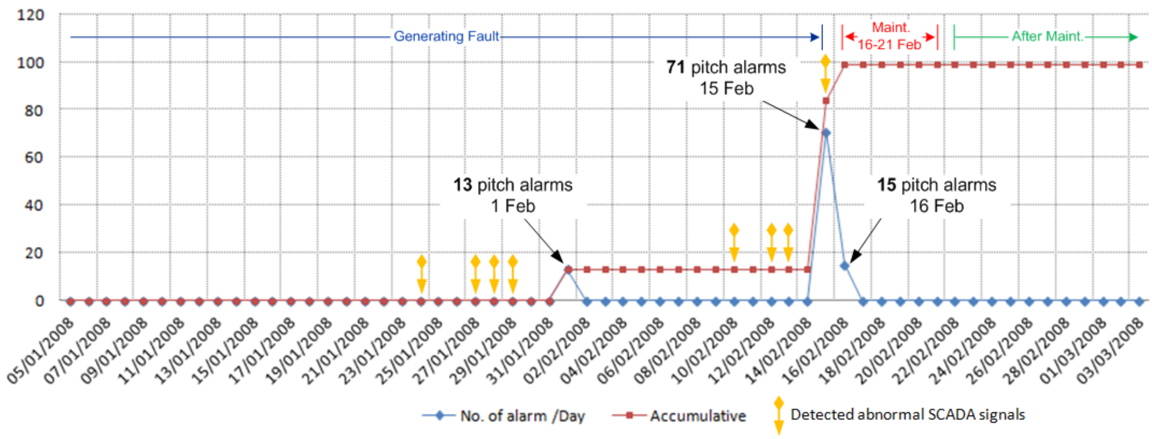


Figure 3: Day-by-day analysis

Considering the economic factors of the WT O&M based on the findings received here, this research is extended to develop an automated fault prognosis for WT pitch system. A diagnosis procedure is proposed in the following section with introducing the ANFIS.

3 Proposed ANFIS Diagnosis Procedure

The fusion of neural networks and fuzzy logic in the ANFIS model provides machine learning as well as readability [15]. Engineers find this useful because the models can be interpreted and supplemented by process operators. With these advantages, an ANFIS diagnosis procedure was proposed, as shown in Figure 4, consisting of the following 4 modules:

- **Data Acquisition:** This module will collect data from the SCADA system and ensure no maintenance or manual stop in the selected period.
- **Feature Extraction:** Data must be not NULL and subject to factory supplied ranges, for example wind speed range from 0m/s to 25m/s. Then, valid data are divided into signals and alarms. Four features, mentioned in Section 2, will be extracted from signals. Alarm

distribution & showers [11] will then be produced from alarm data to validate the final ANFIS result.

- **Multiple ANFIS Diagnosis:** The four features will be passed to the corresponding ANFIS to calculate the fault degree. The overall ANFIS result will be the aggregation of the 4 individual ANFISs, defined as:

$$Result = \frac{\sum_{i=1}^{4} w_i * ANFIS_i}{4}$$

  where w_i is the corresponding weight. All w_i were set to 1 for calculating the average in this case.
- **Fault Diagnosis Result:** Finally, the ANFIS result will be checked against SCADA alarms to provide the warning to the WF operator.

3.1 Training ANFIS

In order to construct the proposed ANFIS diagnosis procedure, the data of the six known pitch faults were used as a knowledge base for training and testing the individual ANFIS.

The fault behaviours of the four features can be represented using a matrix as follows:

$$P_i = [I_{i,1}, I_{i,2}, O_i]^T, \qquad i \in [1,2,3,4]$$

where P_i can be considered to characterise pitch fault. $I_{i,1}$ and $I_{i,2}$ are inputs of the i th feature. The O_i is the corresponding output and it takes one of the values 0 and 1, which indicate the *No* and *Yes* state of the pitch fault. Thus, abnormal data, such as a possible pitch fault, were given value 1 and the remainders were given value 0, to represent *No* pitch fault. By putting six pitch faults' data together, 26,971 sets of data were collected, as shown in Figure 5. In addition, an a-priori knowledge approach was applied, by adding manual created data or using [14], to restrict the value in some specific conditions, as encircled in Figure 5. A hybrid learning strategy was also used, which used a gradient method to update ANFIS premise parameters and a Least Squares Estimate to identify consequent ANFIS parameters [15].

In order to find an optimal structure for each individual ANFIS, batch testing using different numbers of membership function in each input were examined. These calculated the root mean square error of different structures and finally the optimal structures are chose. Then, the dataset were partitioned into two groups: training and testing. Finally, Cases 1-5 provided the training data and Case 6 was used to test the trained model.

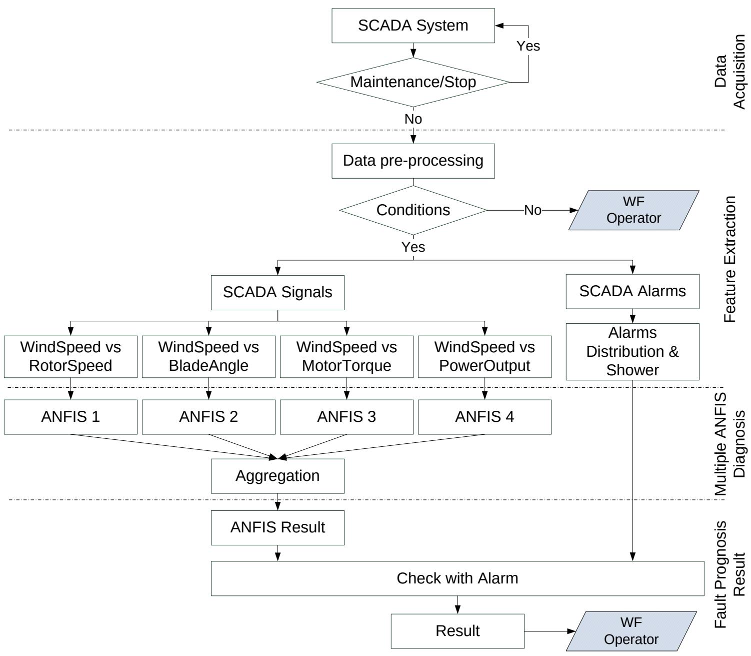


Figure 4: The proposed ANFIS diagnosis procedure.

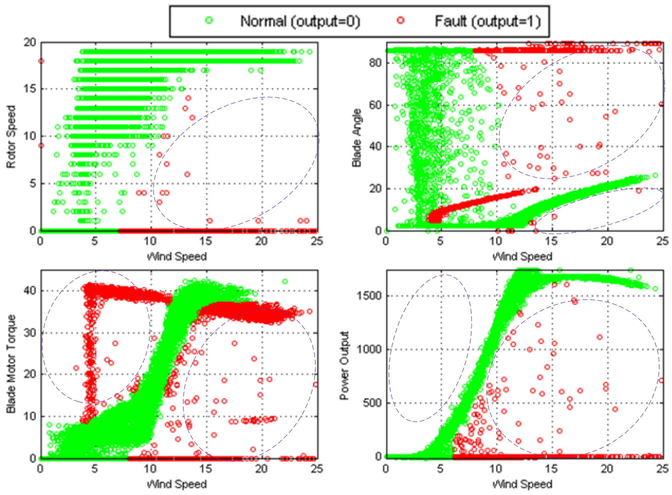


Figure 5: Training Data. Encircled areas have insufficient data and A-priori approach is required.

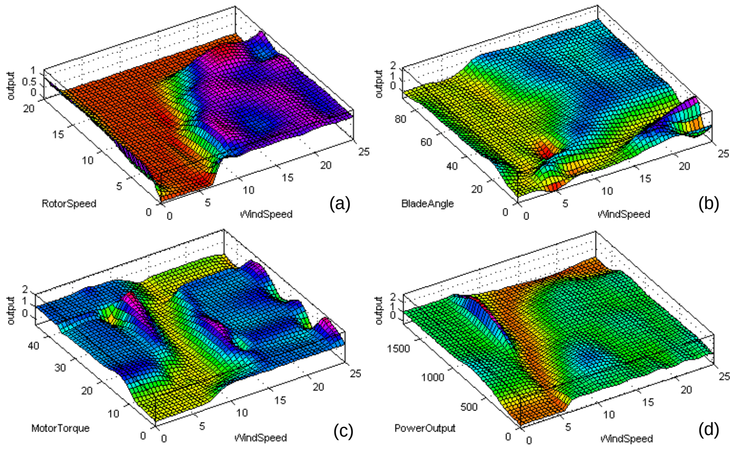


Figure 6: Output surfaces generated from the trained ANFIS

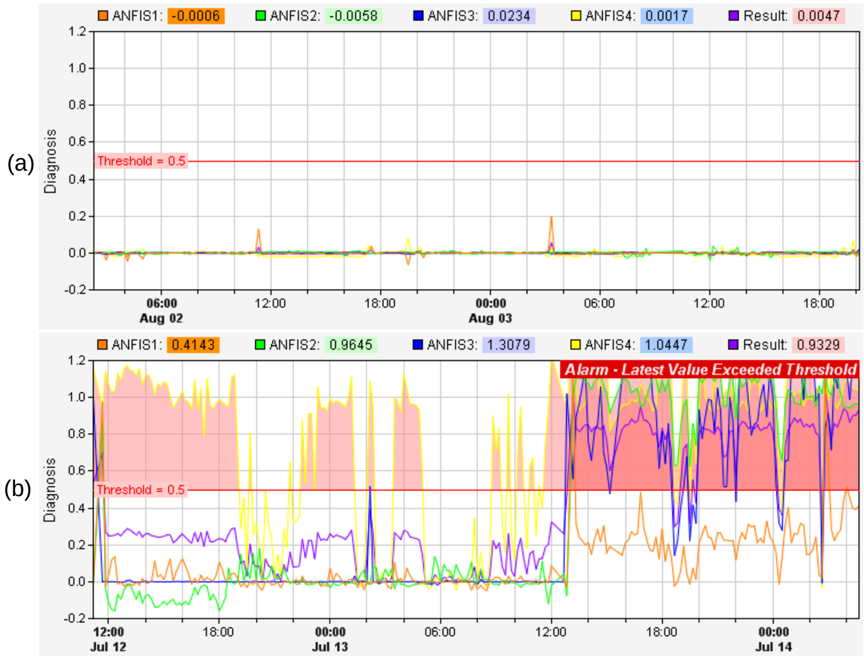


Figure 7: Demonstration of the Diagnosis System (a) A normal running WT; (b) A possible pitch fault has been detected when Threshold is 0.5;

3.2 Trained System

Finally, the output surfaces generated by individual trained ANFIS models are shown in Figure 6. This clearly demonstrates that abnormal data will give a large output, close to 1 as shown in the "Hill", while normal data will give a small output, close to 0 and shown as the "Valley". A demonstration of the ANFIS Diagnosis System is shown in Figure 7, where Figure 7(a) demonstrates a normal running WT and Figure 7(b) demonstrates the detection of a possible pitch fault for which an "Alarm" has been triggered.

4 Validation and Fault Prognosis

4.1 Data for Validation

In order to demonstrate the feasibility of the proposed ANFIS diagnosis procedure, the trained system was tested against the 26 WTs WF, to demonstrate the prognosis of pitch faults. Results were then compared to an Alarm approach to demonstrate the advantage of the prognostic horizon.

The data period was 28 months, from 01/06/2006 to 30/09/2008. For the selected 26 WTs, 910 pitch corrective maintenance records were found in this period, these were further reduced to 487 according to the following 2 criteria.

- Two or more pitch corrective maintenances occurring on the same day;
- A maintenance followed by another maintenance within an interval of not more than 2 days;

Some summary statistics for this WF are:

- An average of 18.7 pitch effective corrective maintenances per WT in this period;
- That is average of 0.67 pitch effective corrective maintenances per WT per month.

4.2 Fault Prognosis using ANFIS Diagnosis Procedure on SCADA Signals

An algorithm was written to apply the trained ANFIS Diagnosis Procedure to calculate the prognostic horizon for every pitch corrective maintenance activity. The Pseudo-code is shown in Table 2. Some usable horizons; 7, 14 and 21 days, were tested to avoid the false identifications. For example a horizon of 180 days would likely to identify a fault that should be independent of current corrective maintenance. In addition, a Threshold and Window Size were required, defined as follows:

- **Threshold (T)** is the aggregation result of the 4 ANFISs, as shown in Fig. 4. Its range is from 0 to 1 for this case.
- **Window Size (WS)** is the number of the consecutive data. Window Size: 6, 18 and 48, were tested to avoid the false identification.

One of the prognosis results is shown in Figure 8. The x-axis is the prognostic horizon in days with y-axis the number of pitches corrective maintenance activities. Each data group is for the proposed window sizes (WS) and thresholds (T). The *unpredicted maintenance* showing in graph legend is the number of unpredicted pitch corrective maintenance activities, out of 487. Figure 8 clearly shows that the proposed ANFIS approach gives the significant warning of pitch faults with a long prognostic horizon up to 21 days.

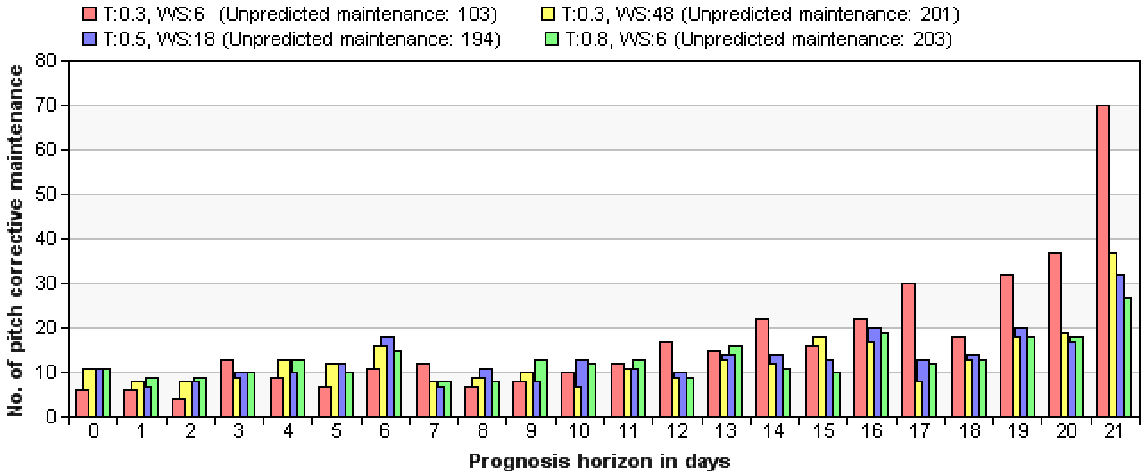


Figure 8: Plot of distribution of SCADA Signals prognosis horizon in days. (T stands for Threshold and WS stands for Window Size).

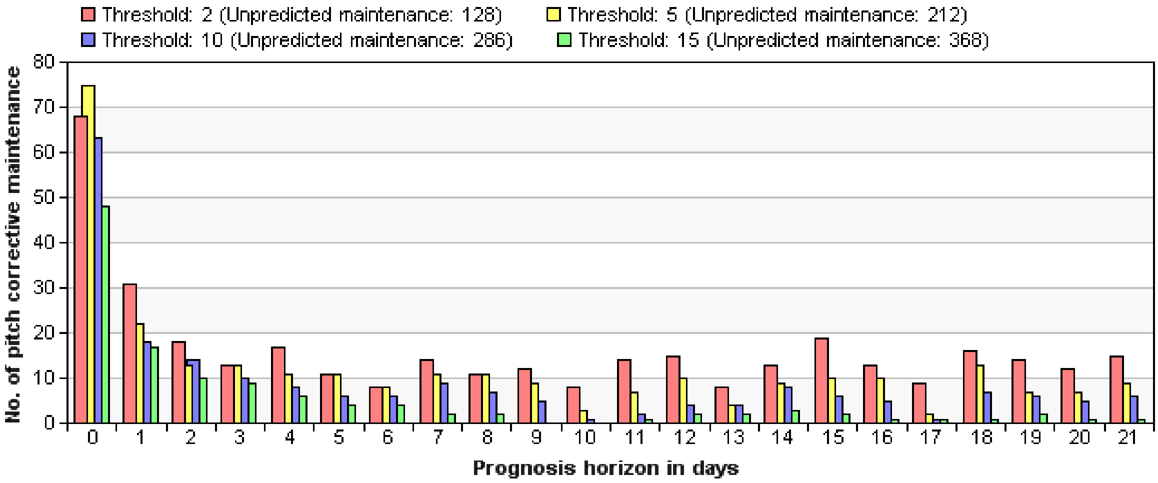


Figure 9: Plot of distribution of SCADA Alarms prognosis horizon in days.

```
Step 1:
Data Cleansing – remove data when it has maintenance;
Step 2:
For each WT in the WF
    For each "pitch corrective maintenance record" in the selected WT
        Within the given usable horizons (7, 14 and 21 days)
            Find the earliest date when Window_Size = (6, 48 or 18) and Threshold >= (0.3, 0.5 or 0.8)
                Prognosis_Day = Maintenance_date – The_Earliest_date

End
```

Table 2: Pseudo-code for calculating the fault prognosis horizon using ANFIS approach

```
Step 1:
Data Cleansing – remove data when it has maintenance;
Step 2:
For each WT in the WF
    For each "pitch corrective maintenance record" in the selected WT
        Within the given usable horizons (7, 14 and 21 days)
            Find the earliest date when total_no_of_pitch_alarm >= Threshold (2, 5, 10 and 15)
                Prognosis_Day = Maintenance_date – The_Earliest_date
End
```

Table 3: Pseudo-code for calculating the fault prognosis horizon using SCADA Alarms

4.3 Fault Prognosis using SCADA Alarms

A common approach to identifying WT faults is counting the number of alarm during a certain period of time. As long as the number of alarms is less than a defined threshold, the situation can be considered safe. In contrast, a possible fault is identified when the number of alarms is larger than the threshold and operators should start to investigate the problem. A study using this approach to examine the efficiency of SCADA pitch alarms for fault prognosis was applied to the same WF to demonstrate the advantage of the proposed ANFIS approach. The threshold was taken as the average SCADA alarms per day. At the beginning of the testing a number of thresholds were considers as follows: 2, 5, 10 and 15. Some usable horizons were tested; 7, 14 and 21 days, to avoid false identifications. A algorithm was also written to calculate the fault prognosis using above approach. The Pseudo-code is shown in Table 3. One of the prognosis results is shown in Figure 9 and we found the Alarm approach gives very little or no prognostic horizon.

5 Result Analysis

Section 4 has demonstrated the proposed approach gives prognostic warning of pitch faults ahead of pitch alarms. In this section, a Confusion Matrix analysis was generated to show the accuracy of the proposed ANFIS approach.

The Confusion Matrix [16] contains information about actual and predicted diagnosis done by the proposed ANFIS system and it is defined as follows:

		Predicted	
		Need Maintenance	No Maintenance
Actual	Had Maintenance	TP	FN
	No Maintenance	FP	TN

- **True Positive (TP):** actual maintenance correctly classified;
- **False Positive (FP):** incorrectly predicted as Needs Maintenance;
- **False Negative (FN):** incorrectly predicted as No Maintenance;
- **True Negative (TN):** all the remainders are correctly classified as No Maintenance;

In addition, a further in-depth analysis of the data is presented utilising:

- Accuracy (Acc) is the proportion of the total number of predictions that are correct.
- Error rate (ER) is the proportion of the total number of predictions that are wrong.
- Recall (RC) is the proportion of actual maintenance cases that are predicted as positive.
- Precision (P) is the proportion of the predicted positive cases that are truly positive.

Defined as follows:

$$ACC = \frac{TP + TN}{TP + FP + TN + FN}$$

$$ER = \frac{FP + FN}{TP + FP + TN + FN}$$

$$RC = \frac{TP}{TP + FN}$$

$$P = \frac{TP}{TP + FP}$$

The Confusion Matrix results of the proposed ANFIS approach applied on the testing WF are shown in Table 4. The table shows the high accuracy and precision of the proposed ANFIS approach. It also can be seen that the precision is increase with the prognostic horizon out to 21 days, whilst the accuracy falls slightly. In addition, recall was improved greatly along with the increase of the prognostic horizon.

	ACC	*ER*	*RC*	*P*
T:0.3 WS:6	88.3%	11.7%	37.0%	76.4%
T:0.3 WS:48	86.0%	14.0%	22.6%	66.1%
T:0.5 WS:18	86.4%	13.6%	21.2%	72.8%
T:0.8 WS:6	86.6%	13.4%	19.6%	79.3%

Usable Prognosis Horizon = 7 days

	ACC	ER	RC	P
T:0.3 WS:6	85.1%	14.9%	48.2%	89.2%
T:0.3 WS:48	80.6%	19.4%	30.7%	83.9%
T:0.5 WS:18	81.0%	19.0%	30.6%	88.5%
T:0.8 WS:6	81.0%	19.0%	29.0%	91.9%

Usable Prognosis Horizon = 14 days

	ACC	ER	RC	P
T:0.3 WS:6	85.9%	14.1%	62.2%	94.4%
T:0.3 WS:48	79.4%	20.6%	43.3%	92.1%
T:0.5 WS:18	79.3%	20.7%	41.8%	94.4%
T:0.8 WS:6	78.9%	21.1%	39.4%	96.2%

Usable Prognosis Horizon = 21 days

Table 4: Confusion matrix results with different usable prognosis horizons

6 Conclusion

From the above results we can draw the following conclusions:

- The proposed ANFIS approach gave significant warning of pitch faults with a prognostic horizon up to 21 days, depending on the ANFIS window size and threshold;
- SCADA alarms also correctly detected pitch faults but counting them gave very little or no prognostic horizon of impending pitch faults;
- Confusion Matrix analysis of the SCADA signal analysis has shown that regardless of window size and threshold the precision of prediction increases with the prognostic horizon out to 21 days, whilst the accuracy falls slightly;
- These results all suggest that whilst SCADA alarm analysis may help to identify pitch fault root causes they cannot predict faults, whereas SCADA signal analysis using ANFIS gives good prediction with a prognostic horizon up to 21 days, a valuable period for WF Operators.

This paper has demonstrated that the proposed ANFIS approach gives prognostic warning of pitch faults ahead of pitch alarms. The SCADA signal analysis using ANFIS has strong potential to provide automated WF fault detection and prognosis. In addition, the proposed ANFIS diagnosis procedure, as shown in Figure 4 & 7, also looks suitable for real-time fault diagnosis.

Acknowledgements

The authors acknowledge the EU FP7 Project RELIAWIND 212966 and UK EPSRC SuperGen Wind EP/H018662/1 Projects. The data used in the paper derived from SCADA data from a number of different WF sources.

References

[1] Feng Y, Tavner PJ, Long H. Early experiences with UK Round 1 offshore wind farms. Proceedings of the Institution of Civil Engineers 2010; 163(Energy No 4): 167–181.

[2] Musial W. Energy from offshore wind, Offshore Technology Conference (OTC2006), Houston, 2006. USA, OTC18355.

[3] EWEA, “Wind in power,” February 2012. http://www.ewea.org/fileadmin/ewea_documents/documents/publications/statistics/Stats_2011.pdf. Last accessed 8th Aug 2012.

[4] P.J. Tavner, J.P. Xiang, F. Spinato, (2006). Reliability analysis for wind turbines, Wind Energy, 10(1), July 2006.

[5] Hameed Z, Hong YS, Cho YM, Ahn SH, Song CK. Condition monitoring and fault detection of wind turbines and related algorithms: a review. Renewable Sustainable Energy Reviews, 2009; 13(1): 1–39. DOI: 10.1016/j.rser.2007.05.008.

[6] Y. Feng, Y.N. Qiu, C.J. Crabtree, H. Long and P.J. Tavner, “Use Of SCADA And CMS Signals For Failure Detection & Diagnosis Of A Wind Turbine Gearbox”, in: EWEA, 2011, Brussels Belgium.

[7] M.A. Sanz-Bobi, M.C. Garcia, P.J. Del. “SIMAP: intelligent system for predictive maintenance application to the health condition monitoring of a wind turbine gearbox”, Computers in Industry 2006; 57: 552–568.

[8] J. Moorse, “Analysis of SCADA Data from Large Wind Farms to Provide

Incipient Fault Detection”, MEng Thesis, August 2010, Durham University.
[9] A. Zaher, S.D.J. McArthur, D.G. Infieeld, Y. Patel, “Online Wind Turbine Fault Detection Through Automated SCADA Data Analysis”. Wind Energy, 2009, 12(6) , 574-593.
[10] B. Chen, Y.N Qiu, Y. Feng. P.J. Tavner and W.W. Song, “Wind Turbine SCADA Alarm Pattern Recognition”, IET RPG Conference 2011.
[11] Y.N. Qiu, Y. Feng, P.J. Tavner, P. Richardson, G. Erdos, B. Chen, “Wind Turbine SCADA Alarm Analysis for Improving Reliability”, Wind Energy.
[12] B. Chen, P.J. Tavner, Y. Feng, W.W. Song, Y.N. Qiu, Bayesian Networks for Wind Turbine Fault Diagnosis, EWEA2012, Poster Presentation, in Copenhegan, Denmark, April 2012.
[13] F. D. Bianchi, H. D. Battista and R. J. Mantz, Wind Turbine Control Systems: Principles, Modelling and Gain Scheduling Design. London: Springer, 2007.
[14] A. Tewari, Prior Knowledge based identification of TSK fuzzy model for static nonlinear systems, PhD Thesis, The Pennsylvania State University, 2009.
[15] J. S. R. Jang, C. T. Sun, and E. Mizutani, Neuro–Fuzzy and Soft Computing, 1997 :Prentice-Hall.
[16] Confusion matrix, http://webdocs.cs.ualberta.ca/~eisner/measures.html, Last accessed 6th Dec 2012.